

This week:

Database :

- Quadtree data structure

I achieve the Quadtree data structure in java. And build the Quadtree of taxi trajectory data.

- Memory management optimization:

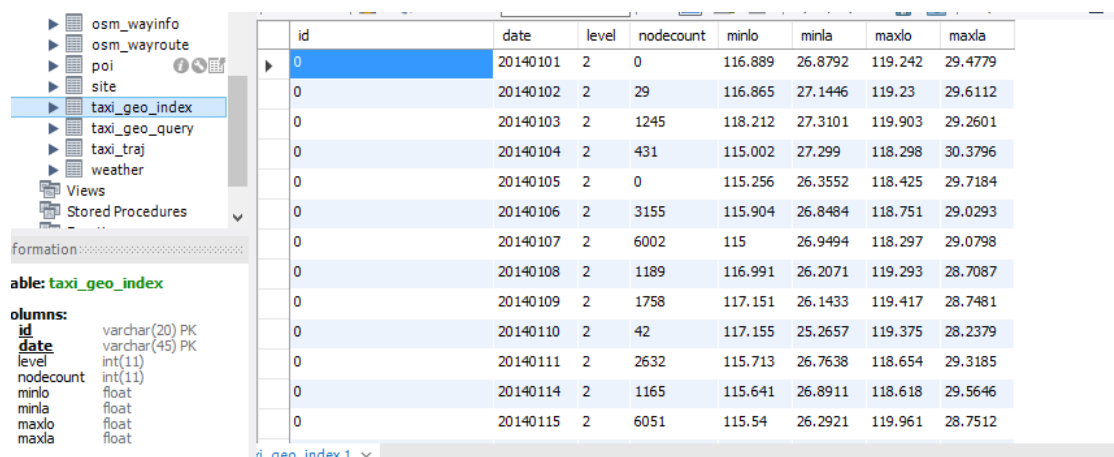
When run the program on taxi trajectory data, it is a problem to manage the memory consumption. The solution is that I build a loop based on one day's data and when the trajectory load in to the memory, it is add to a leaf on our quadtree. As the result, it cost 10min and 10GB for each day's data.

- Some error:

The column name can not be "index"

- Quadtree achieve in mysql

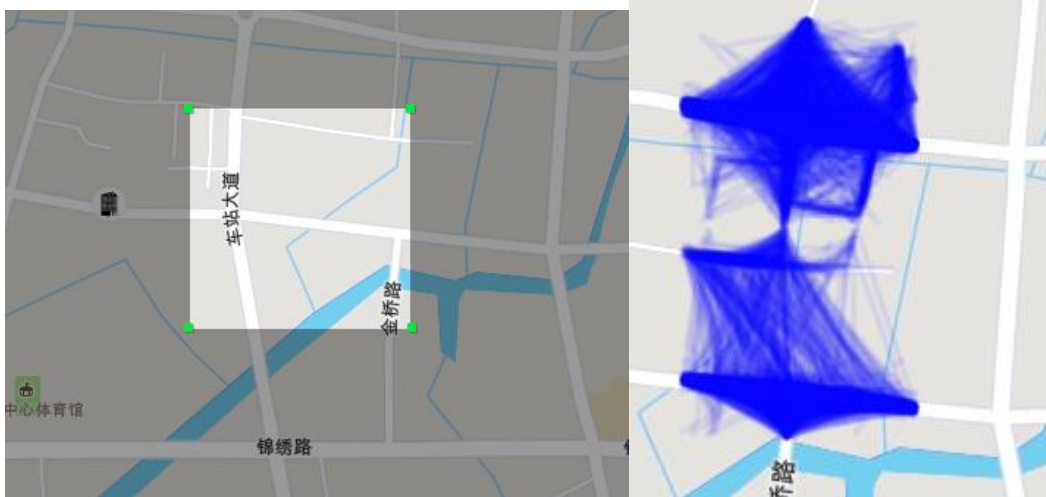
When I achieve the quadtree in java, I put the data into MYSQL. I create two table to build the Quadtree index and store the data. New our query time in entire time area is below 5s (average below 0.5s) which is a few minute before.



id	date	level	nodecount	minlo	minla	maxlo	maxla
0	20140101	2	0	116.889	26.8792	119.242	29.4779
0	20140102	2	29	116.865	27.1446	119.23	29.6112
0	20140103	2	1245	118.212	27.3101	119.903	29.2601
0	20140104	2	431	115.002	27.299	118.298	30.3796
0	20140105	2	0	115.256	26.3552	118.425	29.7184
0	20140106	2	3155	115.904	26.8484	118.751	29.0293
0	20140107	2	6002	115	26.9494	118.297	29.0798
0	20140108	2	1189	116.991	26.2071	119.293	28.7087
0	20140109	2	1758	117.151	26.1433	119.417	28.7481
0	20140110	2	42	117.155	25.2657	119.375	28.2379
0	20140111	2	2632	115.713	26.7638	118.654	29.3185
0	20140114	2	1165	115.641	26.8911	118.618	29.5646
0	20140115	2	6051	115.54	26.2921	119.961	28.7512

System:

- For our database support spatio-temporal query, I achieve a button to support spatio-temporal query in system.



Blog:

I write 5 blog which put out on our group's microblog

Reading:

- The Author-Topic Model for Authors and Documents

文章介绍了 author-topicmodel 的建立方法，主要是 topic，author 和他的区别，应用每一个作者与主题模型上的多项分布，每一个主题和单词上的多项分布，建立混合分布的主题模型。应用在会议文件和作者，关键词的数据集进行了验证。

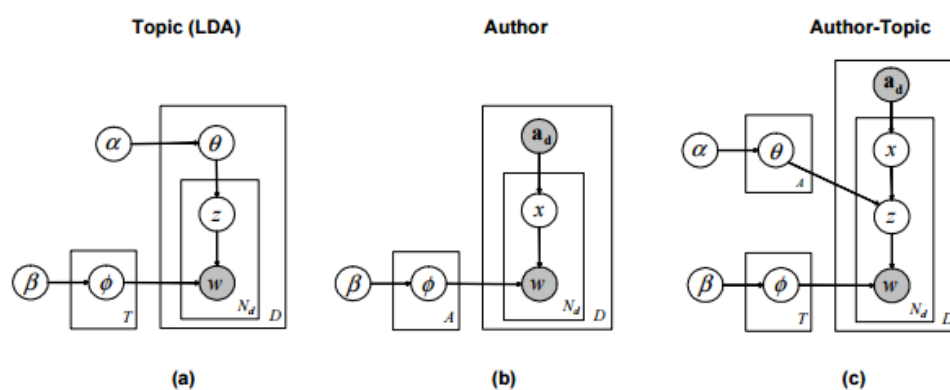


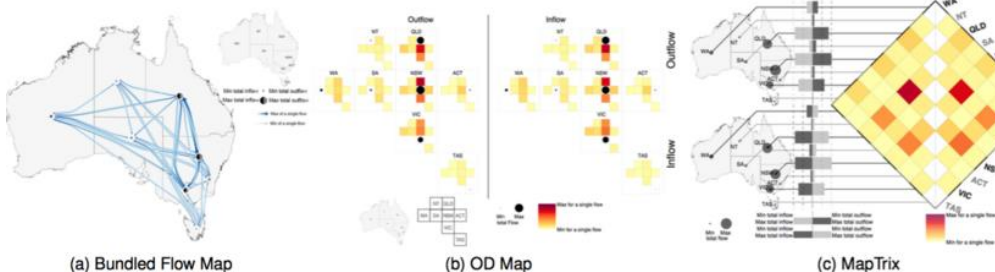
Figure 1: Generative models for documents. (a) Latent Dirichlet Allocation (LDA; Blei et al., 2003), a topic model. (b) An author model. (c) The author-topic model.

- Annotation Graph: A Graph-Based Visualization for Meta-Analysis of Data based on User-Authored Annotations



当人们分析数据的时候，通过对数据进行添加标签，评价等来完成对用户数据特征挖掘的帮助。A, b, c 是探索数据界面，用户可以通过在界面中完成对数据整体的了解，然后选择而某一些轨迹进行进一步的可视分析，可以在 d 视图中通过标签、评价的方式对探索的数据进行注释，然后在 e 视图中展示各个数据注释之间的关系。

● Many-to-Many Geographically-Embedded Flow Visualisation: An Evaluation



我去读了一下这篇文章，本文中的注释添加方法也许可以用于我们的 POI 视图比较。

本周完成了数据的索引，下周也许可以尝试一下。

● Reducing Snapshots to Points: A Visual Analytics Approach to Dynamic Network Exploration

本文的动态网络结构分析，把每一个时间片上的动态网络投影为平面上的一个点。能够

帮助用户分析动态网络随时间变化的情况，我本周让同学做了 PCA 的 POI 投影，下周

可以参考本文做一些别的有意思的尝试，也许某地区的轨迹投影，感觉可能会有好玩的 case。